Title

Optimal designs for asymmetric sigmoidal response curves in bioassays and immunoassays

Authors
Hyun, Seung Won
Wong, Weng Kee
Yang, Yarong

Publication Date
2017-01-17

Peer reviewed

Journal Title
XX(X):2–22

Reprints and permission:
sagepub.co.uk/journalsPermissions.nav
DOI: 10.1177/ToBeAssigned
www.sagepub.com/
SAGE

Optimal designs for asymmetric sigmoidal response curves in bioassays and immunoassays

Seung Won Hyun[1], Weng Kee Wong[2] and Yarong Yang[3]

Abstract

Recent studies show that asymmetric sigmoidal response curves are not uncommon in biomedical studies. For example, the 5-parameter logistic (5PL) model is frequently used to model and analyze responses from bioassays and immunoassays which can be skewed. Various types of optimal experimental designs for 2, 3 and 4-parameter logistic models have been reported but not for the more complicated 5-parameter logistic (5PL) model. Designs currently used for the 5PL model seem ad-hoc with no formal quantitative assessment of their efficiencies. We construct different types of optimal designs for studying various features of the 5PL model and use them to evaluate efficiencies of commonly used designs in bioassays and immunoassays. We also create a user-friendly software package to search for optimal designs tailor-made to user-specified problems for the 5PL model and evaluate robustness properties of the design under a variation of criteria, model forms and mis-specification in the nominal values of the model parameters. Our design strategies can also account for several objectives with varying degrees of importance. As an application, we generate optimal designs for the 5PL model for real studies in immunoassays and bioassays, and show currently used designs are generally inefficient for statistical inference.

Keywords

Approximate Design, Asymmetric Calibration Curves, D-optimal design, Robust Optimal Design, Toxicology

1 Introduction

In toxicology assays, it is often assumed that the response has a symmetric sigmoidal relationship with dose (concentration) and the logistic regression models such as the 3-parameter parameter (3PL) and the 4-parameter logistic (4PL) model are widely used to capture the symmetrical relationship. Recent studies show that asymmetrical response curves are often observed in various bioassays and immunoassays[1–4]. For such asymmetrical sigmoidal response curves, the 3PL and 4PL models are inappropriate. Gottschalk and Dunn[5] showed that the 5-parameter logistic (5PL) model is able to capture the asymmetric relationships adequately and produce dramatically more accurate inference for the assays compared to results using the 3PL or 4PL models.

Statistical inference for bioassays and immunoassays based on the 5PL model is not new. For example, Findlay and Dillard[6] applied the 5PL model to fit the data for ligand binding assays and Feng et al.[7] presented a Bayesian approach to fit the 5PL model using data from an enzyme-linked immunosorbent assay (ELISA). Dawn et al.[8,9] used a modified 5PL model (5PL-1P) to capture the asymmetry in mixture toxicity assessment. Cumberland[10] discussed the choice between the 4PL and 5PL models for estimation purposes; see also model fitting issues using biological data for these models in Davis et al.[11]. Another application of the 5PL model is Gottschalk and Dunn[12], who applied the model for measuring parallelism and relative potency in biological applications.

The design of a scientific study plays a crucial role in the accuracy of the inference that follows. Many of the above studies for the 5PL model used between 5 and 10 evenly spaced design points on the log scale with equal replications. This seems to be the current practice even though there is very little research in the literature to support such design choices. The only design work we were able to find is that from Manukyan and Rosenberger[13], who found locally D-optimal designs for the 5PL model when the response is binary.

The 3, 4 and 5PL models are frequently used to describe sigmoidal response curves and use of a wrong model can produce highly inaccurate or wrong inference. For example, optimal designs for the 3PL model cannot estimate all parameters in the 5PL model and an optimal design for the 5PL model can perform poorly when the 3PL or 4PL model holds. The implemented design

[1]Research Biostatistics, Johnson and Johnson Medical Devices, Irvine, CA 92618, USA
[2]Department of Biostatistics, UCLA School of Public Health, Los Angeles, CA 90095, USA
[3]Department of Statistics, North Dakota State University, Fargo, ND 58102, USA

Corresponding author:
Weng Kee Wong, Department of Biostatistics, UCLA Fielding School of Public Health, Los Angeles, CA
Email: wkwong@ucla.edu

should therefore be robust to mis-specifications of the nominal values, model assumptions and under a variation of criteria.

In practice, there are several objectives in the study and some are more important than others. This calls for a multiple-objective optimal design that can deliver user-specified efficiencies commensurate with the importance of each of the objectives. Such an optimal design is also appropriate when some parameters in the model are more interpretable than others. For instance in the widely used two-parameter Michaelis-Menten model in the biological sciences, the Michaelis-Menten parameter is more interesting than the saturation parameter because it governs how fast the enzyme-substract kinetics reaction velocity is. It follows that the user should devote more resources to estimating the more interesting parameter or parameters so that they are more accurately estimated. Research to date shows a multiple-objective optimal design generally has overall higher efficiencies across all objectives than any of the single objective optimal designs can provide, see for example[14,15]. The former proposed a graphical approach to find dual-objective optimal designs using efficiency plots and, Hyun and Wong[15] gave a step-by-step approach to find 3-objective optimal designs for a nonlinear model. Of course, if the efficiencies sought under all the objectives are too high, a multiple-objective optimal design may not exist.

This paper has two aims. The first is to find various types of optimal designs for the 5PL model and incorporate model uncertainty at the design stage at the same time. Specific designs that seem helpful for bioassays and immunoassays using the 5PL model are optimal designs to accurately estimate (1) the model parameters in the model or (2) a target dose such as the $EC_{50}(ED_{50})$ that results in having one half of test subject having the maximal expected response. Since model uncertainty is always an issue, we also develop optimal designs for discriminating among the 3PL, 4PL and the 5PL models. Additionally, we find multiple-objective optimal designs for the 5PL model and construct a robust D-optimal design that performs well for estimating the model parameters in the 3PL, 4PL or 5PL models regardless which one of them holds.

Our second aim is to facilitate practitioners implement various optimal designs for the 5PL model. To this end, we provide an R package that generate optimal designs discussed in this paper and compare robustness properties of the designs to various departures from the model assumptions. Because the 5PL model is an extension of the 3PL and 4PL models, our functions can also readily find various optimal designs for the 3PL and 4PL models. As an application, we construct an optimal design that performs well for various objectives and model forms in a real immunoassays and bioassays study and report benefits of using our recommended design versus commonly recommended designs for such studies.

Section 2 describes the response curve, interprets the meaning of each parameter in the 5PL model and the Fisher information matrix. Section 3 presents locally optimal designs for studying 3 interesting features under the 5PL model and the robust D-optimal design that performs well for the 3PL, 4PL, and 5PL models for estimating model parameters. In Section 4, we propose an algorithm with R functions to search for all the optimal designs in this paper. In Section 5, we study sensitivities of the locally D-optimal design for the 5PL model to various mis-specifications in the model assumptions. In Section 6, we recommend an optimal design for use in immunoassays and bioassays, and shows it outperforms currently used designs by practitioners. Section 7 contains a conclusion with a summary of our work and future directions.

2 Background

Let X be the user-selected compact design space from which the design points are selected to observe the observations. Let Y_{ij} be the continuous response from the j^{th} replicate at $x_i \in X, j = 1, \ldots, n_i, i = 1, \ldots, K$. Assume that we have resources to take a predetermined number of observations N so that $n_1 + \ldots + n_K = N$. Given a design criterion, the design questions are the optimal number (K) of design points to use, the optimal number of replicates (n_i) and the location of each design point $x_i, i = 1 \ldots, x_K$.

Our statistical models have the form

$$Y_{ij} = f(x_i, \Theta) + \epsilon_{ij} \quad , j = 1, 2, \ldots, n_i, i = 1, 2, \ldots, K, n_1 + \cdots + n_K = N, \quad (1)$$

where $f(x_i, \Theta)$ is the mean response at x_i. The errors ϵ_{ij}s are independent and normally distributed with means 0 and unknown variance σ^2.

We focus on approximate designs or large sample designs where we approximate each n_i/N by its proportion. We denote such a design by $\xi = \{(x_i, w_i)\}_1^K$ where each $x_i \in X,\ w_i \in (0,1)$ and $w_1 + \ldots + w_K = 1$. In dose-response experiments, optimal design issues concern the total number of concentrations to be used (K), where these K concentration levels or design points x_i $x_i, i = 1, 2, \ldots, K$ are, and the proportions $w_i, i = 1, \ldots, K$ of subjects to be allocated to each of these concentrations. Approximate designs can be studied under a unified framework and there are algorithms for finding many types of optimal designs. Formulae for such designs are available for many models and they facilitate studying properties of the optimal approximate designs. In addition, there are theoretical tools for verifying if an approximate design is optimal among all designs and the optimal approximate designs do not depend on the value of N by definition.

We measure the worth of a design by its Fisher information matrix. For the approximate design ξ, the normalized Fisher information matrix is

$$\mathrm{I}(\xi;\Theta) = \frac{1}{\sigma^2}\sum_{i=1}^{K} w_i g(x_i)^\top g(x_i), \tag{2}$$

where $g(x) = \left(\frac{\partial f(x,\Theta)}{\partial \theta_1}, \quad \frac{\partial f(x,\Theta)}{\partial \theta_2}, \quad \ldots, \quad \frac{\partial f(x,\Theta)}{\partial \theta_v} \right)$ and v is the number of model parameters. Since a 'large' information matrix is desirable for statistical inference, many optimality design criteria seek a design that makes this matrix as large as possible in different ways.

For the 5-parameter logistic(5PL) model, the mean response $f(x,\Theta)$ is

$$f(x,\Theta) = \frac{\theta_1 - \theta_4}{\left[1 + (\frac{\theta_3}{x})^{\theta_2}\right]^{\theta_5}} + \theta_4, \tag{3}$$

where θ_1 and θ_4 are the maximum and the minimum expected responses respectively, θ_2 controls the stiffness of the response curve, θ_3 is the position of the transition region in concentration, and θ_5 is the asymmetric factor and positive. the parameters θ_2 and θ_5 jointly control the slope of the response curve. Clearly, the 5PL model becomes the 4-parameter logistic(4PL) model when θ_5 takes the value of 1, and it becomes the 3-parameter logistic(3PL) model when θ_4 and θ_5 take the values of 0 and 1 respectively.

For the 5PL model, the vector $g(x)$ has components given by

$$\begin{aligned}
\frac{\partial f(x,\Theta)}{\partial \theta_1} &= (1+D)^{-\theta_5};\\
\frac{\partial f(x,\Theta)}{\partial \theta_2} &= -\frac{(\theta_1-\theta_4)\theta_5}{\theta_2} D(1+D)^{-1-\theta_5} log(D);\\
\frac{\partial f(x,\Theta)}{\partial \theta_3} &= -\frac{(\theta_1-\theta_4)\theta_2\theta_5}{\theta_3} D(1+D)^{-1-\theta_5};\\
\frac{\partial f(x,\Theta)}{\partial \theta_4} &= 1-(1+D)^{-\theta_5};\\
\frac{\partial f(x,\Theta)}{\partial \theta_5} &= -(\theta_1-\theta_4)(1+D)^{-\theta_5} log(1+D),
\end{aligned}$$

where $D = \left(\frac{\theta_3}{x}\right)^{\theta_2}$. A direct calculation shows the normalized Fisher information matrix (2) is $\mathrm{I}(\xi;\Theta) = A\mathrm{M}(\xi;\Theta)A^\top$, where

$$\mathrm{A} = \begin{pmatrix} 1 & 0 & 0 & 0 & 0 \\ 0 & \frac{-(\theta_1-\theta_4)\theta_5}{\theta_2} & 0 & 0 & 0 \\ 0 & 0 & \frac{-(\theta_1-\theta_4)\theta_2\theta_5}{\theta_3} & 0 & 0 \\ -1 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & -\theta_1+\theta_4 \end{pmatrix},$$

$$\mathrm{M}(\xi;\Theta) = \frac{1}{\sigma^2}\sum_{i=1}^{K} w_i g^*(D_i)^\top g^*(D_i), \quad \varphi(D_i) = (1+D_i)^{-\theta_5}$$

and $g^*(D_i) = \varphi(D_i)\begin{pmatrix} 1, & \frac{D_i log(D_i)}{1+D_i}, & \frac{D_i}{1+D_i}, & \varphi(D_i)^{-1}, & log(1+D_i) \end{pmatrix}$. The matrix A does not contain any concentration level or weight and this implies that maximizing some function of $\mathrm{M}(\xi;\Theta)$ is equivalent to maximizing the same function in $\mathrm{I}(\xi;\Theta)$. We observe that $\mathrm{M}(\xi;\Theta)$ contains only the 3 parameters θ_2, θ_3, and θ_5, and so any classical optimal design such as D-, A-, c-, or D_s-optimal design for model (3) does not depend on the parameters θ_1 and θ_4.

3 Optimal designs

In practice, the choice of an optimal design depends on the objective or objectives of the study. They can vary from estimating all or some model parameters to predicting mean response at a location in the design space or minimizing the sum of elements in the covariance matrix. Frequently, the criterion is formulated as a convex function of the information matrix so that we can use an equivalence theorem to check if the design is optimal among all designs. The equivalence theorem is derived from the directional derivative consideration and is unique for each convex functional. However they all have a general form in terms of an inequality with 0 on the right hand side of the inequality. The function on the left hand side of the inequality is called the sensitivity function of the design in the literature; see design monographs such as Fedorov[16].

The information matrix for a nonlinear model depends on the model parameters and so our optimal design depends on the unknown parameters. Such designs are termed locally optimal[17]. These optimal designs can be sensitive to the nominal values and so they must be selected carefully. However, they are the easiest to find and are important because they typically represent a first step to finding more complex designs[18]. Two approaches that do not require a set of single best guesses for the values of the parameters to find optimal designs are the Bayesian and minimax or maxmin approaches. The former requires a prior distribution for the parameters and the latter requires the user to specify a plausible region of possible values for all the parameters. Both methods generalize the concept of locally optimal designs and both Bayesian and minimax type of optimal designs are clearly much more difficult to find, theoretically or computationally, than locally optimal designs. For example, minimax optimal designs are found under a non-differentiable criterion and there are no effective algorithms for finding them for a general regression model. Chen[19] provides examples of minimax and standardized minimax optimal designs, including showing how locally optimal designs are first determined before a standardized minimax optimal design is found. Recent work on maximin optimal designs, which are equivalent to minimax optimal designs, and Bayesian optimal designs are Coffey[20] and McCallum and Bornkamp[21], respectively, among others. For space consideration, these two approaches will not be further discussed here.

We now review several types of locally optimal designs commonly used in practice. In what is to follow we use the terms locally optimal designs and optimal designs interchangeably when there is no room for confusion.

3.1 D-optimal designs for estimating Θ

D-optimal designs are the most appropriate when the interest in the study is to estimate the vector of model parameters Θ as accurately as possible. The D-optimal design ξ_D maximizes the determinant of the Fisher information matrix $\mathrm{I}_{(\xi;\Theta)}$: $\Psi = |\mathrm{I}_{(\xi;\Theta)}|$. Equivalently, for fixed Θ, we want a design that minimizes the convex function $-log\, |\mathrm{I}_{(\xi;\Theta)}|$. The directional derivative of the D-optimality criterion leads to the sensitivity function: $d_D(x,\xi) = g(x)\mathrm{I}^{-1}(\xi;\Theta)g(x)^\top - 5$. Under the 5PL model, the Equivalence Theorem for D-optimality is that the design ξ_D is the D-optimal design if and only if for all x in X

$$d_D(x,\xi_D) \leq 0,$$

with equality at the design points of the design ξ_D.

3.2 c-optimal designs for estimating the EC_{50}

c-optimal design is used when the goal is to estimate a function of model parameters. The EC_{50} is the concentration producing a response that is half way between the expected maximum and minimum responses of the curve. Under the 5PL, the EC_{50} is expressed as a function of Θ and is given by

$$EC_{50} = \arg_x \left\{ f(x,\Theta) = \frac{1}{2}(\theta_1+\theta_4) \right\} = \theta_3(2^{\theta_5^{-1}} - 1)^{-1/\theta_2}.$$

Let $\widehat{EC}_{50}$ be the maximum likelihood estimate of EC_{50}. The c-optimal design for estimating the EC_{50}, ξ_c minimizes the asymptotic variance of estimating the EC_{50}: $\Psi = Var(\widehat{EC}_{50}) = EC'_{50}\mathrm{I}^{-1}(\xi;\Theta)[EC'_{50}]^\top$, where EC'_{50} is the derivative of the EC_{50} with respect to Θ:

$$EC'_{50} = \begin{pmatrix} 0, & \frac{1}{\theta_2^2}EC_{50}log(2^{\theta_5^{-1}}-1), & \frac{1}{\theta_3}EC_{50}, & 0, & \frac{1}{\theta_2\theta_3^{\theta_2}\theta_5^2}2^{\theta_5^{-1}}EC_{50}^{\theta_2+1}log(2) \end{pmatrix}.$$

Consideration of the directional derivative of the c-optimality criterion leads to the sensitivity function:

$$d_c(x,\xi) = \frac{(g(x)\mathrm{I}^{-1}(\xi;\Theta)[EC'_{50}]^\top)^2}{EC'_{50}\mathrm{I}^{-1}(\xi;\Theta)[EC'_{50}]^\top} - 1,$$

and the Equivalence Theorem states that the design ξ_c is the c-optimal design if and only if for all x in X

$$d_c(x,\xi_c) \leq 0,$$

with equality at the design points of the design ξ_c.

3.3 D_s-optimal designs for estimating the θ_5

D_s-optimal design is used when the goal is to estimate a subset of the model parameters. Due to this characteristic, it is often used for testing nested models. When the 5PL model is used, one interesting question is discriminating between the 4PL and 5PL models (i.e, testing the significance of $\theta_5 = 1$ since the 5PL model becomes the 4PL model when $\theta_5 = 1$). To discriminate between the 4PL and the 5PL model effectively, we need to find a design that minimizes the asymptotic variance of the estimated θ_5 or find the D_s-optimal design for estimating it with minimum variance.

The Fisher information matrix (2) can be partitioned as

$$\mathrm{I}(\xi;\Theta) = \frac{1}{\sigma^2}\begin{pmatrix} \mathrm{I}_{11}(\xi;\Theta) & \mathrm{I}_{12}(\xi;\Theta) \\ \mathrm{I}_{21}(\xi;\Theta) & \mathrm{I}_{22}(\xi;\Theta) \end{pmatrix}, \tag{4}$$

where $\mathrm{I}_{\mathrm{uv}}(\xi;\Theta) = \sum_{i=1}^{K} w_i g_u(x_i)^\top g_v(x_i)$, $u,v = 1,2$. Here

$$g_1(x) = \begin{pmatrix} \frac{\partial f(x,\Theta)}{\partial\theta_1}, & \frac{\partial f(x,\Theta)}{\partial\theta_2}, & \frac{\partial f(x,\Theta)}{\partial\theta_3}, & \frac{\partial f(x,\Theta)}{\partial\theta_4} \end{pmatrix}$$

and

$$g_2(x) = \begin{pmatrix} \frac{\partial f(x,\Theta)}{\partial\theta_5} \end{pmatrix}.$$

The variance of the estimated θ_5 is equal to

$$\mathrm{I}^{22}(\xi;\Theta) = \{\mathrm{I}_{22}(\xi;\Theta) - \mathrm{I}_{21}(\xi;\Theta)\mathrm{I}_{11}^{-1}(\xi;\Theta)\mathrm{I}_{12}(\xi;\Theta)\}^{-1}.$$

The D_s-optimal design ξ_{D_s} for estimating θ_5 is the one that minimizes this variance and equivalently the one that maximizes the determinant

$$\Psi = |\mathrm{I}_{22}(\xi;\Theta) - \mathrm{I}_{21}(\xi;\Theta)\mathrm{I}_{11}^{-1}(\xi;\Theta)\mathrm{I}_{12}(\xi;\Theta)| = \frac{|\mathrm{I}(\xi;\Theta)|}{|\mathrm{I}_{11}(\xi;\Theta)|}.$$

The directional derivative of the D_s-optimality criterion for estimating a subset of s of the parameters leads to the sensitivity function:

$$d_{D_s}(x,\xi) = \left\{g(x)\mathrm{I}^{-1}(\xi;\Theta)g(x)^\top - g_1(x)\mathrm{I}_{11}^{-1}(\xi;\Theta)g_1(x)^\top\right\} - s.$$

In our case, since we have only one parameter of interest, we set $s = 1$. The Equivalence Theorem states that the design ξ_{D_s} is the D_s-optimal design if and only if for all x in X

$$d_{D_s}(x,\xi_{D_s}) \leq 0,$$

with equality at the design points of the design ξ_{D_s}. In this case, D_s-optimal design for estimating θ_5 is equivalent to c-optimal design for estimating θ_5 since the c-optimal design minimizes the asymptotic variance of the estimated θ_5.

3.4 Design efficiency

We use design efficiency to compare the worth of a design relative to the optimum. This is a value between 0 and 1 and frequently it is simply the ratio of the optimal values of the criterion evaluated for the two designs or some simple function thereof. The interpretation of the efficiency of a design is that if its value is r, this design requires $100(1/r - 1)\%$ more observations to do as well as the optimal design. For example, when $e_D(\xi) = 0.5$ for a given design ξ, this design requires 200% more observations to provide the same D-optimality criterion value as the D-optimal design does and this tells that twice as many observations are required for the design to be as efficient as the D-optimal design. The performance of a design ξ for estimating Θ is given by its D-efficiency:

$$e_D(\xi) = \left(\frac{|\mathrm{I}(\xi;\Theta)|}{|\mathrm{I}(\xi_\mathrm{D};\Theta)|}\right)^{\frac{1}{5}}.$$

For estimating a given function of the model parameters, its c-efficiency for estimating EC_{50} is

$$e_c(\xi) = \frac{EC'_{50}\mathrm{I}^{-1}(\xi_\mathrm{c};\Theta)[EC'_{50}]^\top}{EC'_{50}\mathrm{I}^{-1}(\xi;\Theta)[EC'_{50}]^\top},$$

ξ_c is a EC_{50}-optimal design. In particular, its efficiency for estimating the single parameter θ_5, is

$$e_{D_s}(\xi) = \frac{|\mathrm{I}(\xi;\Theta)|/|\mathrm{I}_{11}(\xi;\Theta)|}{|\mathrm{I}(\xi_{\mathrm{D_s}};\Theta)|/|\mathrm{I}_{11}(\xi_{\mathrm{D_s}};\Theta)|}.$$

In practice, we want the implemented design to have high efficiency under the user-specified criterion and preferably a design with high efficiencies under a variation of criteria and a broad class of models.

3.5 A Robust D-optimal design to model mis-specification

Locally optimal designs for nonlinear models can be sensitive to mis-specification in the regression model and the nominal values for the model parameters. The 5PL model becomes the 3PL or 4PL model when some of its model parameters take on specific values. Here we propose a robust D-optimal design that works well for estimating the model parameters for the 3PL, 4PL, and 5PL models.

Let Θ_1, Θ_2, Θ_3 be the vectors of model parameters for the 3PL, 4PL, 5PL models respectively and let $g^1(x)$, $g^2(x)$, $g^3(x)$ be the gradients of the mean functions under the 3 models respectively. The normalized Fisher information matrices for the 3 models are

$$\mathrm{I}_t(\xi;\Theta_t) = \frac{1}{\sigma^2}\sum_{i=1}^{K} w_i g^t(x_i)^\top g^t(x_i),\ t = 1, 2, 3.$$

Following Cook and Wong[14] and Atkinson[22], we use a compound design criterion to construct an efficient design to estimate the model parameters accurately regardless which one of the 3 models holds. The sought design is the one maximizes a user-selected weighted average of the 3 D-optimality criteria for the 3 models:

$$\Psi = \sum_{t=1}^{3} \frac{\lambda_i}{p_t} log(|\mathrm{I}_t(\xi;\Theta_t)|), \tag{5}$$

where for simplicity, we suppress dependence of the criterion on the design, weights and the model parameters. Here, p_t is the number of model parameters for each model, each λ_i is non-negative and $\sum_{t=1}^{3}\lambda_t = 1$. λ_i may be viewed as the prior probability that the ith model holds. By taking directional derivative of the compound criterion, one can show the sensitivity function of the criterion is

$$d_{RoD}(x,\xi) = \left\{\sum_{t=1}^{3} \frac{\lambda_t}{p_t} d_t(x,\xi)\right\} - 1. \tag{6}$$

Here $d_t(x,\xi) = g^t(x)\mathrm{I}_t^{-1}(\xi;\Theta_t)g^t(x)^\top$ and by the Equivalence Theorem, the design ξ_{RoD} is the robust D-optimal design if and only if for all x in X

$$d_{RoD}(x,\xi_{RoD}) \leq 0,$$

with equality at the design points of the design ξ_{RoD}.

4 An algorithm and R-package

Yang[23] introduced an efficient algorithm to search several types of optimal designs for nonlinear models and showed that it outperforms other well known standard design algorithms. Hyun, Wong and Yang[24] modified their algorithm to search the optimal designs more efficiently and this modified algorithm is used to search all the optimal designs in this paper. The modified algorithm runs based on the sensitivity function $d(x,\xi)$ (i.e, the directional derivative of the optimality criterion) and the first and the second derivatives of the optimality criterion with respect to the weights, $\frac{\partial\Psi}{\partial w_i}$ and $\frac{\partial^2\Psi}{\partial w_i w_j}$. The algorithm selects good initial design points via the Fedorov's algorithm[16] and at each iteration, it selects a design point that maximizes $d(x,\xi)$ and adds to the previously selected design points. Then the optimal weights for the selected design points are obtained by the Newton Raphson's method using $\frac{\partial\Psi}{\partial w_i}$ and $\frac{\partial^2\Psi}{\partial w_i w_j}$. Lastly the obtained optimal design is verified by the General Equivalence Theorem[25]. Base on the algorithm, we developed an R package **Opt5PL** to search all the optimal designs in this paper and the package is available in the Comprehensive R Archive Network (`https://CRAN.R-project.org/package=Opt5PL`). The description showing examples of searching optimal designs using the package is given in the supporting web materials of this paper. This R package contains several functions (`ROPT`, `EDpOPT`, `DsOPT`, `Deff`, `EDpeff`, `Dseff`) useful for obtaining and evaluating optimal designs under the 5PL models. Some functions also can be used to study optimal designs for the 3PL or 4PL models.

The `ROPT` function obtains the robust D-optimal design for the model uncertainty between the 3PL, 4PL, and 5PL models. In addition, the function can be used to obtain D-optimal design for each of the 3 models. The `ROPT` function maximizes the compound optimality criterion (5) in this paper and verifies the optimality of the generated design using the General Equivalence Theorem by producing a graphical plot of the sensitivity function (6).

The `EDpOPT` and `DsOPT` functions obtain c-optimal design for estimating the EC_p and D_s-optimal design design for estimating θ_5 under the 5PL model respectively, and they also verify the obtained optimal designs using the the General Equivalence Theorem. Here the EC_p is the concentration level that achieves the $100p\%$ of the difference between the maximum and the minimum responses. By setting $p = 0.5$, the `EDpOPT` function can find the c-optimal design for estimating the EC_{50} under the 5PL model.

The `Deff`, `EDpeff`, `Dseff` functions obtain D-efficiency for estimating Θ, c-efficiency for estimating the EC_p, and D_s-efficiency for estimating θ_5 under the 5PL model respectively. These functions obtain the efficiencies of any given design ξ. The function `Deff` also can be used to compute the D-efficiencies under the 3PL and 4PL models. In the following applications, all the optimal designs for the 5PL model and their performances are obtained using this R functions.

5 Robustness of D-optimal design for the 5PL model

5.1 Robustness to the model parameter values and to the 4PL model

The 5PL model can capture more various dose-response relationships compared to the 4PL model since it can control an asymmetric levels of the response curve. We believe there must be some advantages of using the optimal designs for the 5PL model over the optimal designs for the 4PL model. Here we compare the D-optimal design for the 5PL model to practically efficient optimal designs for the 4PL model and check the robustness of the D-optimal design to various model parameter values and to the different model, the 4PL model.

Holland[26] presents practically useful cD-optimal designs for estimating the model parameters and the EC_{50} simultaneously under the 4PL model. The cD-optimal design combines the requirements for the c- and D-optimal designs in some ways, so that the obtained design provides a user-specified efficiency for estimation of both the model parameters and the EC_{50}. These optimal designs work well for the 4PL model that describes symmetrical sigmoidal response curves. However, there is a doubt if the optimal designs under the 4PL model still work well for the 5PL model which can account for asymmetrical relationships.

Here we consider their 3 cD-optimal designs that provide competitive efficiencies for the two objectives under the 4PL model and ascertain whether they still perform well for the same objectives under the 5PL model. The 3 optimal designs ξ_i^{4PL}, $i = 1, 2, 3$ were obtained using 3 approaches: i) maximizing a weighted geometric average of c- and D- criteria; ii) two stage procedure using the idea of design augmentation in Padmanabhan[27]; iii) maximizing a weighted geometric average of c- and D- criteria under some constraint. So the 3 designs correspond to the designs 2,3b, and 4b in Table 1 of Holand-Latz's paper[26]. We compare them with the D-optimal design for the 5PL model ξ_D^{5PL} and evaluate the robustness of the 4 designs to mis-specified parameter values.

The same setup in Holland[26] is used: a true value of $\Theta_2 = (1, 1, 1, 0)$ for the 4PL model and set $\theta_5 = 1$ for the 5PL model, so $\Theta_3 = (1, 1, 1, 0, 1)$; and the log concentration range is between -5 and 5. Table 1 shows the performances of the 4 designs ξ_1^{4PL}, ξ_2^{4PL}, ξ_3^{4PL}, and ξ_D^{5PL} for estimating the model parameters and the EC_{50} under the 5PL model using D- and c-efficiency, e_D^{5PL} and e_c^{5PL}. To ascertain whether the optimal designs are robust to nominal values of the model parameter, we fix the true value of $\theta_5 = 1$ and compute optimal designs with values of θ_5 different from unity and compare changes in their efficiencies when we make various assumptions on the asymmetrical aspects of the response curves. Table 1 shows these efficiencies when nominal values for θ_2 and θ_3 are fixed. Table A1 and A2 in Web Appendix A show corresponding results for different values of θ_2 and θ_3.

Table 1 about here

The cD-optimal designs for the 4PL model perform poorly overall for the 5PL models with various asymmetric factor levels. The design ξ_2^{4PL} performs better than the other two cD-optimal designs but their efficiencies are unsatisfactory and they change dramatically in the c-efficiency for different values of θ_5. Similar results are observed for the cD-optimal designs for different values of θ_2 and θ_3 under the 5PL model. This is an expected result because the cD-optimal designs are obtained under the 4PL model and optimal designs for nonlinear models usually become inefficient when they are used for a different model. One interesting finding is the cD-optimal designs don't perform well even when $\theta_5 = 1$ for the 5PL model. The 5PL model becomes identical to the 4PL model when $\theta_5 = 1$ but their Fisher information matrices are still different, so the optimal designs for the 4PL model becomes inefficient when they are used for the 5PL model. This reinforces the importance of considering the 5PL model to construct the optimal designs and shows how the designs obtained from the 4PL model perform inefficiently when they are used for the 5PL model which can account for the asymmetrical relationship.

In contrast, the D-optimal design ξ_D^{5PL} works well for estimating the model parameters for the different values of θ_2, θ_3, and θ_5. The c-efficiencies of ξ_D^{5PL} are lower than its D-efficiencies but mostly they are higher than ones for the cD-optimal designs, and they don't show any dramatic changes except the cases when the values of θ_2 and θ_5 are far from the true value such as $\theta_2, \theta_5 = 0.5$ or 2.0. It seems the D-optimal design for the 5PL model is more resistant to the changes in the values of θ_3 than the changes in θ_2 and θ_5. In a similar way, we also check the robustness of ξ_D^{5PL} to the 4PL model. Under the 4PL model, θ_3 represents the EC_{50} and θ_2 represents the slope at the EC_{50}. Table 2 assumes the 4PL model and compares the D-efficiencies, e_D^{4PL}, and the c-efficiencies, e_c^{4PL} of the 4 designs for different values of θ_3. Table A3 in Web Appendix A in the supporting web materials displays results for different values of θ_2.

Table 2 about here

The D-optimal design for the 5PL model ξ_D^{5PL} performs well for both objectives under the 4PL model for the various values of θ_3. For different values of θ_2, ξ_D^{5PL} still provides competitive D- and c-efficiencies compared to the cD-optimal designs but the efficiencies becomes lower when the slope becomes higher($\theta_2 > 1.8$). Since ξ_D^{5PL} does not take into account c-optimality criterion for estimating the EC_{50}, it provides relatively lower c-efficiencies than D-efficiencies for both the 4PL and the 5PL model but the c-efficiencies can be improved by applying the same approaches to combine c- and D-optimality criteria using the 5PL model.

5.2 Robustness to the 5PL model for binary response data

In this paper, the 5PL model is used to define the mean of a continuous response. Sometimes the 5PL model is used to define the probability of observing an interesting event(i.e., P(Y=1)) for a binary response. When the 5PL model is used for a binary response, the Fisher information matrix is written as $\mathrm{I_B}(\xi;\Theta) = \frac{1}{P(1-P)}\sum_{i=1}^{K} w_i g(x_i)^\top g(x_i)$, where $P = f(x,\Theta)$ and $g(x)$ is still the gradient of the 5PL model. For a binary response, the information matrix is different from one for a continuous response $\mathrm{I}(\xi;\Theta)$ in (2), so the optimal designs under the 5PL model for a binary response become different from ones for a continuous response. For simplification, from now on, the 5PL model for a continuous response and the 5PL model for a binary response are denoted by $5PL^C$ and $5PL^B$ respectively. In this section, we want to see how the optimal designs for $5PL^C$ work under $5PL^B$.

Here the same log concentration range $[-5,5]$ and the same standard parameterization for Θ_3 are used to obtain the optimal designs. In order to check the robustness of the optimal designs for $5PL^C$ to $5PL^B$, D-, c- and D_s-efficiencies based on $\mathrm{I_B}(\xi;\Theta)$, $e_D^{5PL^B}$,$e_c^{5PL^B}$, and $e_{D_s}^{5PL^B}$ are computed, and we see how the D-, c- and D_s-optimal designs for $5PL^C$ works compared to the corresponding optimal designs for $5PL^B$ under various parameter values of θ_2, θ_3, and θ_5. Let ξ_D^* denote the D-optimal design for $5PL^C$ with $\Theta_3 = (1,1,1,0,1)$. We also check how the D-optimal design for $5PL^C$ with the standard parameterization works robustly to different objectives and different parameter values under $5PL^B$. Table 3 shows the efficiencies of optimal designs for $5PL^C$ when they are used for $5PL^B$ with various θ_5 values. The same tables for different values of θ_2 and θ_3 are given in the supporting web materials (Tables A4 and A5 in Web Appendix A). It seems the D-optimal design for $5PL^C$ performs good when it is used for D-optimality under $5PL^B$ and also it works robustly to different parameter values. In contrast, for c- and D_s- optimality cases, the optimal designs for $5PL^C$ doesn't work effectively when they are used for $5PL^B$.

Table 3 about here

6 Applications

In this section, two real studies are considered to implement the proposed optimal designs and compare with currently used designs by practitioners.

Study 1: Bio-Plex cytokine assays are described extensively in `www.bio-rad.com`, `www.biocompare.com`, among other web sites. We used the setup described in the technical report[11] and considered Bio-Plex cytokine assays that are bead-based multiplex sandwich immunoassays. The models of interest are the 4PL and the 5PL models, which have been shown to be appropriate for fitting data from such assays. There are two recommended setups for the assays in order to achieve efficient performance. One is a high-sensitivity range standards (0.2-3,200 pg/ml) and the other one is a broad range standards (1.95-32,000 pg/ml). The broad range standards is considered to study optimal designs. When the logistic model is considered, at least 5 standards(concentrations) for the 4PL model and 6 standards for the 5 PL model are required but a total of 8 evenly distributed standards in the range are recommended for an accurate fit. Based on a 4 fold dilution series, the broad range standard has 8 design points (1.95, 7.8, 31.25, 125, 500, 2,000, 8,000, 32,000).

Study 2: Dawn et al[8] assessed toxicity of four chemical agents alone and in mixture using the 5PL-1P model. The 5PL-1P model is a modified 5PL model by removing the minimum response parameter. For concentration-response curves, they fitted the model to each single chemical or their mixture at three different exposure durations: 15, 30 and 45 min. The experimental design in the study prepared test concentrations by serial dilution using 1.867 as the dilution factor. Among the four agents, two agents that has the same concentration range(7-300 mg/L) are considered to study optimal designs. One is ethyl chloroacetate(ECAC) and the other one is 3-methyl-2-butanone(3M2B). Based on the dilution factor, the design has 7 design points (7.09, 13.24, 24.71, 46.12, 86.10, 160.70, 300.00) and two replications at each design point.

In what is to follow, we obtain various optimal designs for the 5PL model and evaluate how well the above recommended designs are relative to the optimal designs. The optimal designs are locally optimal in that they depend on the nominal set of values for the model parameters. Specifically, we compare performances of the original designs used in the actual studies to the optimal designs for the 3 criteria discussed earlier. For each study, the original design has equal weights for the given design points.

6.1 Optimal designs for fitting the 5PL model

In Study 1, the investigators used the 5PL model to study cytokine assays over a pre-specified range of concentrations between 1.95 and 32,000. The set of nominal values for the model parameters is $\Theta_3 = (30,000, 1,800, 0.5, 1)$. To simulate various response curves over the same range, we created 6 different sets of possible values of (θ_2, θ_5) commensurate with values of θ_1, θ_3, and θ_4 in their work. Results from previous section suggest that the locally D-optimal design for the $5PL$ model is more sensitive to this two parameters θ_2 and θ_5, and we wish to assess how other designs perform for estimating different aspects of the curve. Figure 1 shows these 6 response curves from the broad range design ξ_{S1}.

Each single chemical in Study 2 has three different response curves based on the three different exposure times. In Dawn's paper[8], they provided the estimated parameter values of the EC_{50}, the slope, and the asymmetric factor when the 5PL-1P model was fitted to each of the agents. So, the 6 different sets of parameter values for the two agents ECAC and 3M2B are adopted to create 6 additional response curves of the 5PL model. The paper didn't provide the parameter estimates for the maximum and the minimum responses. The optimal designs for the 5PL model do not depend on these two parameter values as mentioned earlier. For illustration purpose, we assume their values are 100 and 0 respectively since the response is a toxicity effect(0-100%) in the study. Figure 2 shows these 6 additional response curves from the design in Study 2 ξ_{S2}.

Figure 1 about here

Figure 2 about here

Table 4 shows all the parameter values of the 5PL model for Study 1 and 2. Table A6 and A7 in Web Appendix A in the supporting web material lists the locally D-optimal designs ξ_D, locally c-optimal designs ξ_c, and locally D_s-optimal designs ξ_{D_s} for the 5PL models. The suggested guideline for fitting the 5PL model requires at least 6 design points, and so one may add two evenly spaced design points between the second and the fourth design points of ξ_D on the log scale, and call this an extended D-optimal design ξ_{ExD}. The extended D-optimal design has equal weight across the seven design points. The two additional design points of ξ_{ExD} are also given in the tables along with other optimal designs.

Table 4 about here

Table 5 shows how the 3 designs, ξ_D, ξ_{ExD}, and the original design ξ_{S1} or ξ_{S2} perform under the 3 objectives. For both studies, it seems that the locally

D-optimal design still has high efficiencies for the other two objectives. For estimating the EC_{50} in Study 1, it provides various efficiencies between 0.55 and 0.90. It seems the efficiency becomes lower when the slope of the response curve is stiff but the D-optimal design provides satisfactory efficiency for estimating the EC_{50} when the response is gradually increased. The extended D-optimal design cannot perform well as the D-optimal design does but it still provides efficiencies of between 0.60 and 0.70 regardless of the different parameter values. The actual design ξ_{S1} in Study 1 works poorly for estimating the EC_{50} and the model discrimination and it provides nearly 0 efficiencies when the slope is steep. The actual design ξ_{S2} in Study 2 performs close to the extended D-optimal design for all three objectives. If researchers want to find optimal designs that provide more balanced efficiencies or user-specified efficiencies for the 3 objectives, the multiple-objective optimal design approach in Hyun, Wong and Yang[24] can be applied. The overall message from the table is that the extended D-optimal design will be practically useful since it provides constant efficiencies disregarding the changes in the parameter values, and it performs much better than the actual design when the response curve has a stiff slope. In addition, it works as well as the D-optimal design when the goal is for estimating the model parameters and the efficiencies for the other objectives can be increased by adopting the multiple-objective optimal design approach.

Table 5 about here

6.2 Robustness of the optimal designs to different models

Most commonly used models in immunoassays and bioassays are the 4PL and the 5PL models and sometimes the 3PL model. All these 3 models are used to describe a sigmoidal curve for the mean response. Now we want to find a design that works well for the 3 models. The robust D-optimal design introduced in this paper can provide well balanced efficiencies for estimating model parameters under the 3 models.

The model parameter values for the 3 models need to be specified in order to find the robust D-optimal design. For illustration, we assume that the 6 sets of parameter values for each of Study 1 and 2 are true for the 5PL model and the mean responses of the 5PL model at the design points of the actual design are used to estimate the model parameters, $\hat{\Theta}_1$ and $\hat{\Theta}_2$ for the 3PL and the 4PL models respectively. For Study 1, eight mean responses corresponding to the eight design points of ξ_{S1} are generated from the 5PL model with the given parameter values and use them to obtain the least square estimates of the parameters for the 3PL and the 4PL models, and the parameter estimates for the 3PL and the 4PL models for Study 2 are obtained in the same way using ξ_{S2}.

Their estimates are obtained using *nlm* in *R* program. Based on the criterion (5) and assuming the 3 models are equally plausible (i.e., $\lambda_1 = \lambda_2 = \lambda_3 = 1/3$), the robust D-optimal designs ξ_{RoD} are obtained for both studies (see Table A8 and A9 in the Web Appendix A). The table shows that the robust D-optimal design always includes the 2 endpoints, and the middle points and their weights are changed variously.

Table 6 shows the D-efficiencies of the designs ξ_{S1}, ξ_{S2}, ξ_{ExD}, ξ_D, ξ_{RoD} across the different models (3PL,4PL,5PL) under the 6 different sets of parameter values for Study 1 and 2. The D-efficiencies under the 3PL and 4PL models are calculated using the definition in Section 3.4 assuming one of the 3 models is the true model and the parameters are correctly specified. For Study 1, the actual design has reasonable efficiencies under the 4PL and 5PL models when the response curves are not too stiff but the D-optimal design and the extended D-optimal design provide higher efficiencies all the time. For Study 2, the actual design performs very well compared to the D-optimal design and the extended D-optimal design. The designs, ξ_{S1}, ξ_{S2}, ξ_{ExD}, and ξ_D become less efficient when they are used for the 3PL model. In contrast, our proposed compromised design ξ_{RoD} is quite robust to the model uncertainty since it provides much higher efficiencies for the 3PL model while maintaining quite competitive efficiencies for the 4PL and 5PL models.

Table 6 about here

7 Conclusions

We present optimal designs for studying several meaningful features of the 5PL model, which has been shown it can fit bioassays data better than 3 or 4PL models when the response curve is asymmetric. We compare performance of the actual designs that are recommended for immunoassays and bioassays and this paper provides quantitative evidence that the assays can be studied more effectively using optimal designs discussed here. To facilitate users implement optimal designs for the 5PL model, we provide an R package **Opt5PL** to assess performance of the actual designs relative to the optimum, and also to study robustness properties of the optimal design to various model assumptions. In particular, we show that the locally D-optimal design for estimating the model parameters in the 5PL model is relatively robust to mis-specified parameter values for θ_2,θ_3 and θ_5 and also to the form of the mean response. However, the same robustness cannot be guaranteed when it is used for different objectives. Additionally, we show more realistic designs, such as multiple-objective optimal designs, should be used to capture the goals of the study more accurately. These designs tend to generally provide noticeably higher efficiencies for all objectives than the single-objective optimal designs.

Our work is the first to address a variety of design issues for the 5PL model. Additionally, we provide an R package to facilitate researchers in bioassays and immunoassays to design more efficient experiments for the 3,4 and 5PL models. Our focus here is on constructing locally optimal designs and future directions for research include finding different types optimal designs using the maximin, Bayesian and multistage approaches. Another interesting design issue not discussed here is finding optimal designs for the 5PL model when the data has heterogeneous variances. Sometimes bioassays data have heterogeneous variances as the concentration changes. It is not known whether optimal designs discussed here are robust to heteroscedastic errors in the model or whether use of optimal designs based on other efficient estimators such as the maximum quasi likelihood estimator(MqLE) or the extended quasi likelihood estimator(EQL) is a better option.

Acknowledgement

Wong was partially supported by a grant from the National Institute of General Medical Sciences of the National Institutes of Health under Award Number R01GM107639. The content is solely the responsibility of the authors and does not necessarily represent the official views of the National Institutes of Health.

Supplementary material

The following supporting information is available as part of the Supplementary material:

Table A1 and Table A2. D-efficiencies, e_D^{5PL}, and c-efficiencies, e_c^{5PL}, of 4 designs under the 5PL model with various values of θ_2 and θ_3.

Table A3. D- and c-efficiencies, e_D^{4PL} and e_c^{4PL} of the 4 designs under the 4PL model with various values of θ_2(slope).

Table A4 and Table A5. The efficiencies of optimal designs for $5PL^C$ when they are used for $5PL^B$ with various θ_2 and θ_3 values.

Table A6 and Table A7. The locally D-, extended D-, c-, D_s-optimal designs for the 5PL model in Study 1 and 2.

Table A8 and Table A9. The Robust D-optimal designs for the 5PL model in Study 1 and 2.

Descriptions for an R package Opt5PL

References

1. Law B. *Immunoassay: A Practical Guide*. CRC Press, 2002. ISBN 9780748405602.
2. DeSilva B, Smith W, Weiner R, Kelly M, Smolec J, Lee B, Khan M, Tacey R, HIll H and Celniker A. Recommendations for the Bioanalytical Method Validation of Ligand-binding Assays to Support Pharmacokinetic Assessments of Macromolecules. *Pharm Res* 2003; 20(11): 1885-1900.
3. Liao JJZ, Duan F, Meng Y and Zhao, Q. Selecting an Appropriate Dose-Response Curve in Bioassay Development. *Frontiers in Drug Design and Discovery* 2010; 5: 67-96.
4. Leschik JJ, Diana T, Olivo PD, Konig J, Krahn U, Li Y, Kanitz M and Kahaly GJ. Analytical Performance and Clinical Utility of a Bioassay for Thyroid-Stimulating Immunoglobulins. *Am J Clin Pathol* 2013; 139: 192-200.
5. Gottschalk PG and Dunn JR. The five-parameter logistic: A characterization and comparison with the four-parameter logistic. *Anal Biochem* 2005a; 343: 54-65.
6. Findlay JW and Dillard RF. Appropriate Calibration Curve Fitting in Ligand Binding Assays. *The AAPS journal* 2007; 9: 260-267.
7. Feng F, Sales AP and Kepler TB. A Bayesian approach for estimating calibration curves and unknown concentrations in immunoassays. *Bioinformatics* 2011; 27: 707-712.
8. Dawson DA, Mooneyham T, Jeyaratnam J, Schultz TW and Poch G. Mixture Toxicity of S_N2-Reactive Soft Electrophiles: 2Evaluation of Mixtures Containing Ethyl a-Halogenated Acetates. *Archives of Environmental Contamination and Toxicology* 2011; 61: 547557.
9. Dawson DA, Genco N, Bensinger HM, Guinn D, IIgiovine ZJ, Wayne Schultz T and Poch G. Evaluation of an asymmetry parameter for curve-fitting in single chemical and mixture toxicity assessment. *Toxicology* 2012; 26: 156-161.
10. Cumberland WN, Fong Y, Yu X, Defawe O, Frahm N and De Rose S. Nonlinear calibration model choice between the four and five parameter logistic models. *J Biopharm Stat* 2015; 25: 972-983.
11. Davis D, Zhang A, Torrence J, et al. Selection of Standards for Bio-Plex Cytokine Assays. Bio-Rad 2900, `http://www.bio-rad.com/LifeScience/pdf/Bulletin_2900.pdf`.
12. Gottschalk PG and Dunn JR. Measuring parallelism, linearity, and relative potency in bioassay and immunoassay data. *J Biopharm Stat* 2005b; 15: 437-463.
13. Manukyan Z and Rosenberger WF. D-optimal design for a five-parameter logistic model. *mODa 9 Advances in Model-Oriented Design and Analysis* 2010: 113-120.
14. Cook RD and Wong WK. On the Equivalence of Constrained and Compound Optimal Designs. *J Am Stat Assoc* 1994; 89: 687-692.

15. Hyun SW and Wong WK. Multiple objective optimal designs to study the interesting features in a dose-response relationship. *Int J Biostat* 2015; 11: 253-271.
16. Fedorov V. *Theory of Optimal Experiments*. Trans. and eds. W.J. Studden and E.M. Klimko Academic. New York, 1972.
17. Chernoff H. Locally optimal designs for estimating parameters. *ANN OF MATH STAT* 1953; 24: 586-602.
18. Ford I, Kitsos CP, Titterington DM. Recent advances in nonlinear experimental design. *Technometrics* 1989; 31: 49-60.
19. Chen RB, Chang SP, Wang W, Tung HC and Wong WK. Minimax optimal designs via particle swarm optimization methods. *Stat Comput* 2015; 25: 975-988.
20. Coffey T. Bioassay case study applying the maximin D-optimal design algorithm to the four-parameter logistic model. *Pharm Stat* 2015; 14: 427-432.
21. McCallum E and Bornkamp B. Accounting for parameter uncertainty in two-stage designs for Phase II dose-response studies. *Modern Adaptive Randomized Clinical Trials.* CRC press, 2015, pp. 427-450.
22. Atkinson AC, Donev AN, Tobias RD. *Optimum experimental designs with SAS*. Oxford: Oxford University Press, 2007.
23. Yang M, Biedermann S, Tang E. On optimal designs for nonlinear models: a general and efficient alg orithm. *J Am Stat Assoc* 2013; 108: 1411-1420.
24. Hyun SW, Wong WK, Yang Y. VNM: An R Package for Finding Multiple-Objective Optimal Designs for the 4-Parameter Logistic Model. *J Stat Softw* 2018; 83: DOI: 10.18637/jss.v083.i05.
25. Kiefer J. *Jack Carl Kiefer Collected Papers III, Design of Experiments*. Springer-Verlay: New York, 2014.
26. Holland-Letz T. On the Combination of c- and D-Optimal Designs: General Approaches and Applications in Dose-Response Studies. *Biometrics* 2017; 73: 206-213.
27. Padmanabhan SK and Dragalin V. Adaptive Dc-optimal designs for dose finding based on a continuous efficacy endpoint. *Biom J* 2010; 52: 836-852.

Table 1. D-efficiencies, e_D^{5PL}, and c-efficiencies, e_c^{5PL}, of 4 designs under the 5PL model with various values of θ_5. The designs $\xi_1^{4PL}, \xi_2^{4PL}, \xi_3^{4PL}$ in the first 3 rows are the cD-optimal designs found from the 3 ways assuming the 4PL model with nominal values $\Theta_2 = (1, 1, 1, 0)$ holds. The fourth design ξ_D^{5PL} is the D-optimal design for the 5PL model constructed from nominal values $\Theta_3 = (1, 1, 1, 0, 1)$. Zero efficiencies mean actual efficiencies are smaller than 0.01.

Design	ξ	θ_5	0.5	0.8	1.0†	1.2	1.5	1.8	2.0
ξ_1^{4PL}	$\begin{pmatrix} -5.00 & -0.60 & -0.50 & 0.50 & 0.60 & 5.00 \\ 0.25 & 0.19 & 0.06 & 0.06 & 0.19 & 0.25 \end{pmatrix}$	e_D^{5PL}	0.23	0.27	0.29	0.30	0.31	0.31	0.31
		e_c^{5PL}	0.00	0.01	0.13	0.46	0.07	0.01	0.01
ξ_2^{4PL}	$\begin{pmatrix} -5.00 & -1.00 & 0.00 & 1.00 & 5.00 \\ 0.25 & 0.12 & 0.25 & 0.12 & 0.25 \end{pmatrix}$	e_D^{5PL}	0.58	0.67	0.71	0.74	0.75	0.74	0.72
		e_c^{5PL}	0.05	0.52	0.97	0.90	0.69	0.50	0.40
ξ_3^{4PL}	$\begin{pmatrix} -5.00 & -0.60 & -0.50 & 0.50 & 0.60 & 5.00 \\ 0.25 & 0.21 & 0.04 & 0.04 & 0.21 & 0.25 \end{pmatrix}$	e_D^{5PL}	0.22	0.25	0.27	0.28	0.29	0.29	0.29
		e_c^{5PL}	0.00	0.01	0.11	0.45	0.06	0.01	0.01
ξ_D^{5PL}	$\begin{pmatrix} -5.00 & -1.96 & -0.15 & 1.65 & 5.00 \\ 0.20 & 0.20 & 0.20 & 0.20 & 0.20 \end{pmatrix}$	e_D^{5PL}	0.86	0.95	1.00	0.97	0.92	0.83	0.76
		e_c^{5PL}	0.37	0.71	0.69	0.64	0.61	0.56	0.53

The true value of θ_5

Table 2. D- and c-efficiencies, e_D^{4PL} and e_c^{4PL} of the 4 designs under the 4PL model with various values of $\theta_3(ED_{50})$. ξ_1^{4PL}, ξ_2^{4PL}, ξ_3^{4PL} are the cD-optimal designs for the 4PL model with $\Theta_2 = (1, 1, 1, 0)$ and ξ_D^{5PL} is the D-optimal design for the 5PL model with $\Theta_3 = (1, 1, 1, 0, 1)$.

Design	θ_3	0.5	0.8	1.0†	1.2	1.5	1.8	2.0
ξ_1^{4PL}	e_D^{4PL}	0.81	0.89	0.90	0.89	0.86	0.83	0.81
	e_c^{4PL}	0.54	0.81	0.84	0.82	0.73	0.62	0.54
ξ_2^{4PL}	e_D^{4PL}	0.84	0.90	0.90	0.90	0.88	0.86	0.85
	e_c^{4PL}	0.56	0.78	0.81	0.80	0.71	0.62	0.56
ξ_3^{4PL}	e_D^{4PL}	0.81	0.89	0.90	0.89	0.87	0.84	0.81
	e_c^{4PL}	0.54	0.81	0.84	0.82	0.74	0.62	0.54
ξ_D^{5PL}	e_D^{4PL}	0.89	0.89	0.90	0.91	0.91	0.91	0.91
	e_c^{4PL}	0.54	0.55	0.55	0.55	0.54	0.55	0.55

The true value of θ_3

Table 3. D-, c-, D_s-efficiencies ($e_D^{5PL^B}$, $e_c^{5PL^B}$, and $e_{D_s}^{5PL^B}$) of the optimal designs for $5PL^C$ when they are used for $5PL^B$ with various values of θ_5. Here $5PL^C$ and $5PL^B$ represents the 5PL models with a continuous response and a binary response respectively. $\xi_D^{\theta_5}$, $\xi_c^{\theta_5}$, and $\xi_{D_s}^{\theta_5}$ are the sought D-, c-, and D_s- optimal designs under $5PL^C$ with the given θ_5 value in the column and ξ_D^* is the sought D-optimal design under $5PL^C$ with $\Theta_3 = (1, 1, 1, 0, 1)$.

Efficiency	θ_5	0.5	0.8	1.0†	1.2	1.5	1.8	2.0
$e_D^{5PL^B}$	$\xi_D^{\theta_5}$	0.82	0.83	0.81	0.80	0.78	0.76	0.73
	ξ_D^*	0.80	0.81	0.81	0.82	0.82	0.81	0.79
$e_c^{5PL^B}$	$\xi_c^{\theta_5}$	0.48	0.67	0.65	0.59	0.29	0.59	0.57
	ξ_D^*	0.22	0.59	0.45	0.36	0.17	0.34	0.34
$e_{D_s}^{5PL^B}$	$\xi_{D_s}^{\theta_5}$	0.54	0.55	0.50	0.46	0.40	0.36	0.34
	ξ_D^*	0.29	0.31	0.32	0.34	0.36	0.34	0.31

The true value of θ_5

Table 4. The 6 sets of parameter values of the 5PL model for Study 1 and Study 2.

Case	Study 1	Study2
c_1	$\Theta = (30000, 0.5, 800, 0.5, 2.0)$	$\Theta = (100, 0.81, 40.14, 0, 1.63)$
c_2	$\Theta = (30000, 0.5, 800, 0.5, 5.0)$	$\Theta = (100, 0.93, 49.82, 0, 1.06)$
c_3	$\Theta = (30000, 1.0, 800, 0.5, 1.0)$	$\Theta = (100, 1.11, 69.26, 0, 0.59)$
c_4	$\Theta = (30000, 1.0, 800, 0.5, 1.5)$	$\Theta = (100, 0.80, 10.58, 0, 2.33)$
c_5	$\Theta = (30000, 2.0, 800, 0.5, 2.0)$	$\Theta = (100, 0.80, 12.12, 0, 2.33)$
c_6	$\Theta = (30000, 2.0, 800, 0.5, 5.0)$	$\Theta = (100, 0.83, 16.93, 0, 1.90)$

Table 5. Efficiencies of the designs for various objectives of the 5PL model under the 6 sets of nominal values for each of Study 1 and 2. The locally D-optimal design, the locally extended D-optimal design and the actual designs, are given, respectively by ξ_D, ξ_{ExD}, and ξ_{S1} and ξ_{S2}.

Study	Efficiency	ξ	c_1	c_2	c_3	c_4	c_5	c_6
1	e_D	ξ_{ExD}	0.91	0.91	0.91	0.91	0.91	0.91
		ξ_{S1}	0.88	0.74	0.86	0.83	0.45	0.32
	e_c	ξ_D	0.82	0.85	0.71	0.66	0.55	0.57
		ξ_{ExD}	0.62	0.67	0.67	0.60	0.63	0.64
		ξ_{S1}	0.55	0.30	0.55	0.47	0.03	0.35
	e_{D_s}	ξ_D	0.84	0.83	0.86	0.85	0.86	0.85
		ξ_{ExD}	0.67	0.66	0.65	0.65	0.65	0.65
		ξ_{S1}	0.59	0.35	0.56	0.48	0.05	0.04
Study	Efficiency	ξ	c_1	c_2	c_3	c_4	c_5	c_6
2	e_D	ξ_{ExD}	0.91	0.91	0.91	0.91	0.91	0.91
		ξ_{S2}	0.92	0.92	0.92	0.90	0.91	0.91
	e_c	ξ_D	0.88	0.90	0.88	0.89	0.90	0.90
		ξ_{ExD}	0.67	0.68	0.67	0.69	0.69	0.69
		ξ_{S2}	0.68	0.68	0.66	0.63	0.64	0.66
	e_{D_s}	ξ_D	0.84	0.85	0.85	0.85	0.84	0.85
		ξ_{ExD}	0.68	0.68	0.68	0.68	0.68	0.68
		ξ_{S2}	0.66	0.66	0.66	0.61	0.62	0.63

Table 6. D-efficiencies of the 4 designs, ξ_{S1} and ξ_{S2}, ξ_{ExD}, ξ_D, and ξ_{RoD} for the 3PL, 4PL and 5PL models under the 6 sets of nominal values for each of Study 1 and 2. ξ_{RoD} is the robust D-optimal design with $\lambda_1 = \lambda_2 = \lambda_3 = 1/3$ and e_D^{3PL}, e_D^{4PL}, e_D^{5PL} are, respectively, the efficiencies of the given design for estimating the model parameters in the 3PL, 4PL, and 5PL models.

Study 1									
Θ	ξ	e_D^{3PL}	e_D^{4PL}	e_D^{5PL}	Θ	ξ	e_D^{3PL}	e_D^{4PL}	e_D^{5PL}
c_1	ξ_{S1}	0.60	0.83	0.88	c_4	ξ_{S1}	0.51	0.81	0.83
	ξ_{ExD}	0.69	0.85	0.91		ξ_{ExD}	0.70	0.86	0.91
	ξ_D	0.69	0.91	1.00		ξ_D	0.68	0.91	1.00
	ξ_{RoD}	0.80	0.94	0.94		ξ_{RoD}	0.79	0.93	0.95
c_2	ξ_{S1}	0.40	0.73	0.74	c_5	ξ_{S1}	0.33	0.66	0.45
	ξ_{ExD}	0.67	0.84	0.91		ξ_{ExD}	0.70	0.87	0.92
	ξ_D	0.67	0.91	1.00		ξ_D	0.66	0.90	1.00
	ξ_{RoD}	0.80	0.94	0.94		ξ_{RoD}	0.80	0.94	0.93
c_3	ξ_{S1}	0.58	0.83	0.86	c_6	ξ_{S1}	0.20	0.46	0.32
	ξ_{ExD}	0.71	0.86	0.91		ξ_{ExD}	0.71	0.87	0.92
	ξ_D	0.69	0.90	1.00		ξ_D	0.66	0.90	1.00
	ξ_{RoD}	0.80	0.93	0.95		ξ_{RoD}	0.80	0.94	0.94
Study 2									
Θ	ξ	e_D^{3PL}	e_D^{4PL}	e_D^{5PL}	Θ	ξ	e_D^{3PL}	e_D^{4PL}	e_D^{5PL}
c_1	ξ_{S2}	0.84	0.86	0.92	c_4	ξ_{S2}	0.86	0.86	0.90
	ξ_{ExD}	0.82	0.84	0.91		ξ_{ExD}	0.84	0.84	0.91
	ξ_D	0.85	0.90	1.00		ξ_D	0.88	0.90	1.00
	ξ_{RoD}	0.90	0.94	0.97		ξ_{RoD}	0.92	0.94	0.97
c_2	ξ_{S2}	0.85	0.86	0.92	c_5	ξ_{S2}	0.86	0.86	0.91
	ξ_{ExD}	0.83	0.84	0.91		ξ_{ExD}	0.84	0.84	0.91
	ξ_D	0.88	0.90	1.00		ξ_D	0.87	0.90	1.00
	ξ_{RoD}	0.92	0.93	0.97		ξ_{RoD}	0.92	0.94	0.97
c_3	ξ_{S2}	0.83	0.86	0.92	c_6	ξ_{S2}	0.85	0.86	0.91
	ξ_{ExD}	0.82	0.84	0.91		ξ_{ExD}	0.84	0.84	0.91
	ξ_D	0.87	0.90	1.00		ξ_D	0.87	0.90	1.00
	ξ_{RoD}	0.91	0.94	0.97		ξ_{RoD}	0.92	0.94	0.97

Figure 1. Response curves for Study 1. The mean responses at design points in ξ_{S1} are plotted.

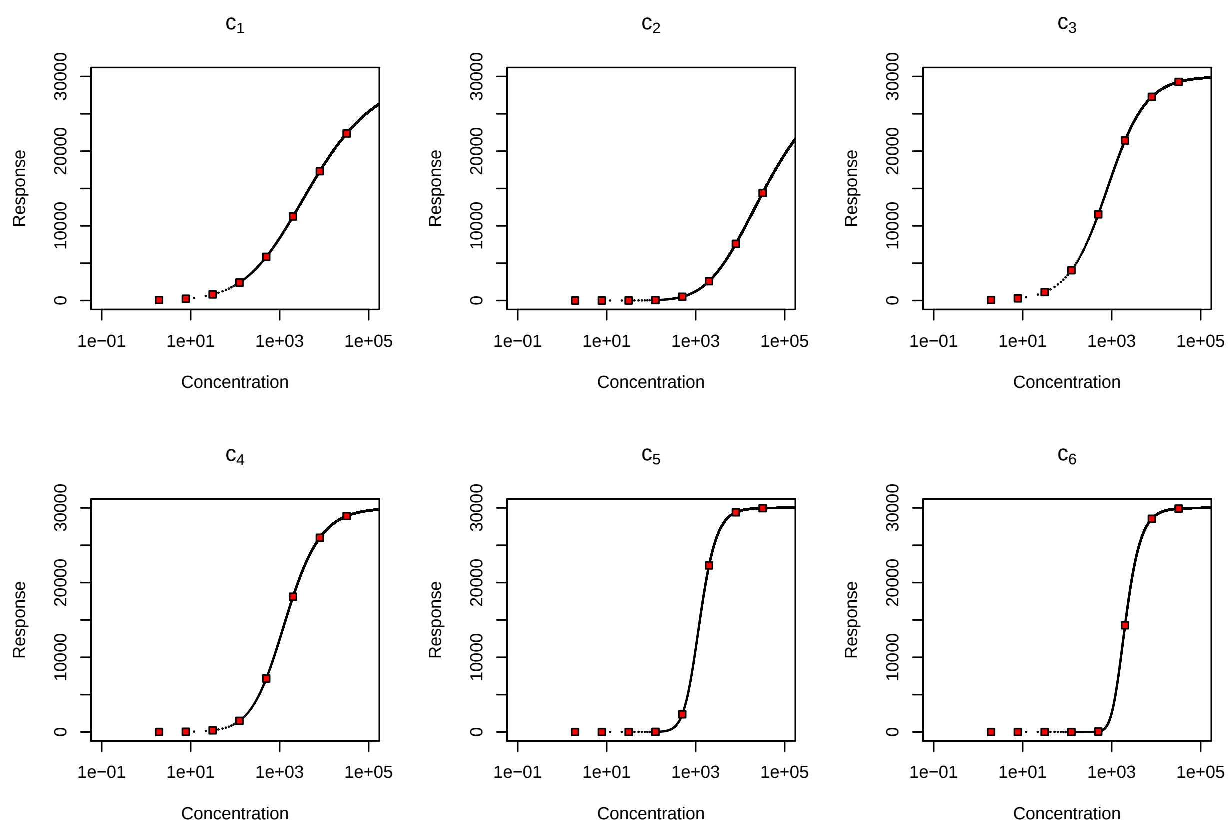

Figure 2. Response curves for Study 2. The mean responses at design points in ξ_{S2} are plotted.